Bringing Comics Into The Digital Age

Bahar Aydemir[1], Deblina Bhattacharjee[1,2], Baran Ozaydin[1], Mathieu Salzmann[2], Sabine Süsstrunk[1]

[1]Image and Visual Representation Lab (IVRL), [2]Computer Vision Lab (CVLAB), EPFL

Goal

Comics (bande dessinée or "BD" in French) have been published in various formats over the years in newspapers, magazines, stand-alone albums, and more recently, digital form. Transferring them from one format to another is costly and time-consuming for artists. This project aims to simplify this process.

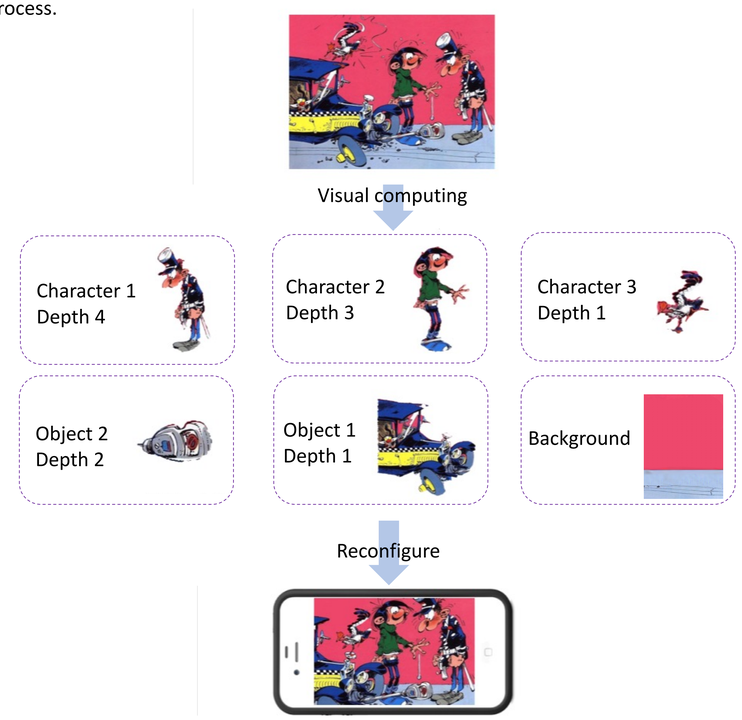
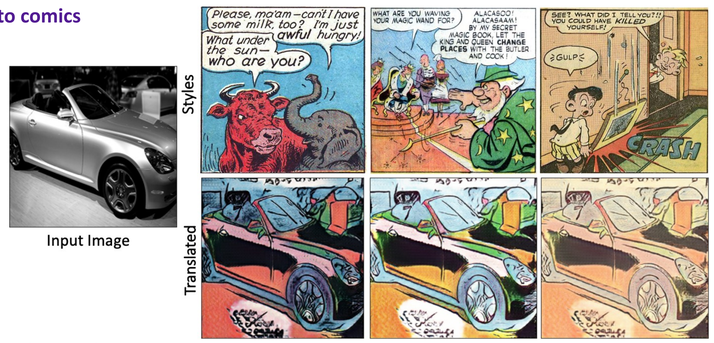


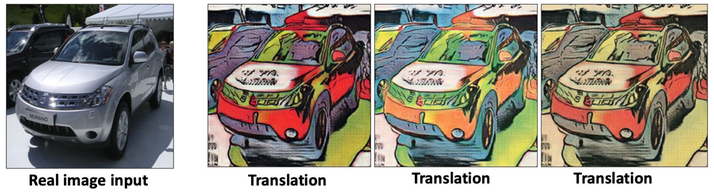
Image to image translation

Image to image translation converts one image into another while preserving the desired characteristics such as image content, objects and their spatial properties.

Real to comics

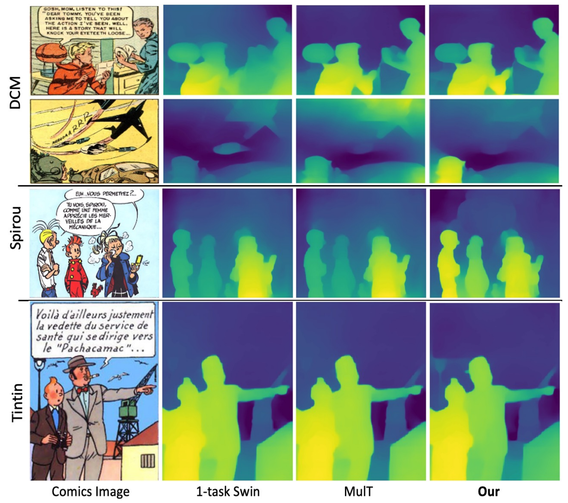


Multistyle translation



Depth Estimation

In depth estimation, we predict the distance between the objects in a single scene or multiple scenes from a given viewpoint.



Challenges in the comics domain

In comics, estimating depth is essential to rearrange the objects in the scene. The depth of the objects are vastly different than the ones in the natural images since comics do not obey the rules of the physical world. The artist can alter the physical constraints as they wish.

Segmentation

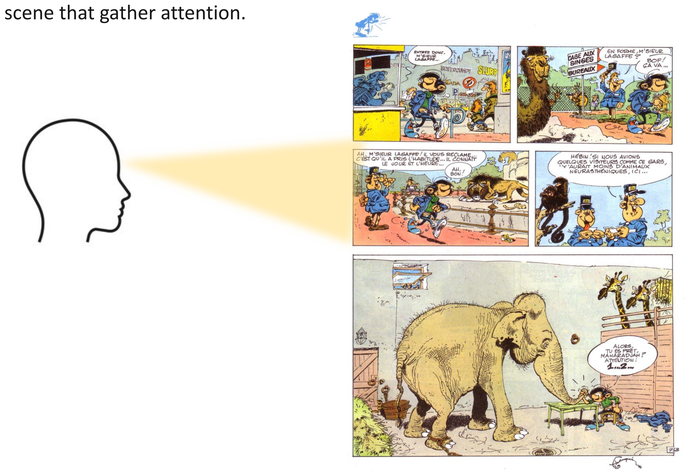
Segmentation is the task of identifying the classes to which the pixels of an image belong.

Challenges in the comics domain

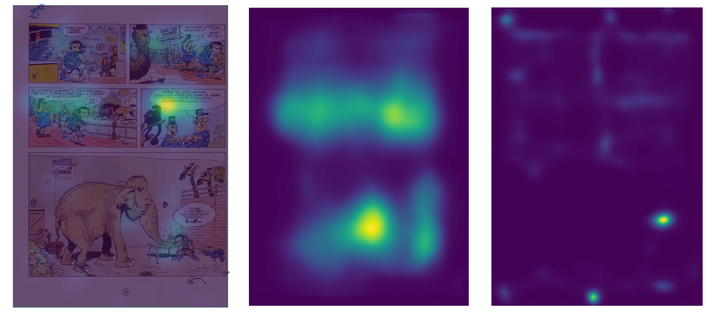
We need segmentation to identify the key elements for reconfiguration. Pretrained networks perform poorly when applied to comics because they have no notion of the artistic styles seen in comics.

Saliency

Visual saliency prediction aims to identify the most significant parts of a scene that gather attention.



Input image



Ground truth collected with eye-tracking

Prediction with the CLIP encoder

Prediction with the VGG encoder

Challenges in the comics domain

Saliency depends on both low level features (color, contrast, brightness) and high level features (semantics, objects). Image to image translation cannot be used for saliency estimation due to this complexity. Lack of annotated data also hinders training models from scratch in comics domain.

References

[1] Deblina Bhattacharjee, Tong Zhang, Sabine Süsstrunk and Mathieu Salzmann, MulT: An End-to-End Multitask Learning Transformer, CVPR, 2022

[2] Bahar Aydemir, Ludo Hoffstetter, Tong Zhang, Mathieu Salzmann and Sabine Süsstrunk, TempSAL: Uncovering Temporal Information For Saliency Prediction, CVPR 2023.

Contact Information

Web : https://ivrl.epfl.ch/

Email: {bahar.aydemir, deblina.bhattacharjee, baran.ozaydin}@epfl.ch